DOCUMENT RESUME

ED 040 714 JC 700 182

AUTHOR Tucker, John M.; Wilkerson, George J.
TITLE Measuring Faculty Effectiveness in the Community College: A Proposal.
PUB DATE [70]
NOTE 16p.

EDRS PRICE EDRS Price MF-$0.25 HC-$0.90
DESCRIPTORS *Behavioral Objectives, *Effective Teaching, *Evaluation, *Evaluation Criteria, Instruction, *Junior Colleges, Personnel Evaluation, Teachers

ABSTRACT

Recent attempts by the Federal Government, industry, and community groups to concern themselves with school accountability suggest that unless the educational community begins to develop effective and meaningful evaluative criteria, external agencies may do it for them. This paper describes the current status of the evaluation of teaching effectiveness, and suggests guidelines for developing a more comprehensive evaluative program. To begin with, a criterion-referenced approach to evaluation is suggested, with greater emphasis placed on the product rather than on the process of teaching. On this basis, changes in learner behavior are seen as the ultimate or most important measurement criterion. After discussing recent efforts to establish effective evaluation schemes and the obstacles which these schemes must overcome, an approach to the problem is outlined. This approach calls for a commitment by faculty and administrators to develop behavioral measures of the individual instructor's and school's effectiveness; the development of applicable pre- and post-test measures of effectiveness; program implementation; behavioral definition of skills necessary for success in various occupations and professions; and finally, the combination of course, curricular, and institutional objectives into a general set of goals for which both instructor and administrator can be held accountable. A brief description of efforts at John Tyler Community College (Virginia) toward these goals concludes the discussion. (JO)

MEASURING FACULTY EFFECTIVENESS IN THE COMMUNITY COLLEGE:

A PROPOSAL

BY

JOHN M. TUCKER

EDUCATIONAL DEVELOPMENT OFFICER
JOHN TYLER COMMUNITY COLLEGE

and

GEORGE J. WILKERSON

CHAIRMAN, DIVISION OF COMMUNICATIONS AND SOCIAL SCIENCES
JOHN TYLER COMMUNITY COLLEGE

THE CURRENT STATUS OF FACULTY EVALUATION

Unless the academic world starts resolving its evaluative problems, others may do it for them. And if this happens these external agencies will exert more control than the schools on the direction and emphasis of education. The problem, therefore, is for those of us in the education community to develop fair, valid, and meaningful criteria for our evaluation before it is done for us.

In the opinion of most writers concerned with faculty evaluation, the current situation is far from satisfactory. Arthur M. Cohen has asserted that "the entire history of faculty evaluation approximates the sordid!" And Cohen and Brawer make a strong case for abandoning all current practices of faculty evaluation (4).

However, evaluation is inevitable. Dressel (5) writes that there is no real issue regarding the presence or absence of evaluation. He says that whenever one is faced with a choice, evaluation, whether conscious or not, is present. Dressel warns that failure to systematically engage in evaluation in reaching the many decisions necessary in education means that decision by prejudice, by tradition, or by rationalization is the result. He asserts that "...such patterns of decision making are not consistent with the aims of (higher) education, which in our culture are based upon the assumption that informed judgments can and should be wiser judgments."

Furthermore, signs of new external pressures from the federal government, industry, and the general public indicate that the time has now come for the academic world to accept the responsibility for resolving the problems involved in evaluation. Otherwise, others will take over the

function of setting up the criteria of evaluation and the academic world will lose even more control over its own destiny.

The federal government is showing increasing concern for visible evidence of the effective use of the millions of dollars that it spends on educational efforts.

The USOE has established new education posts of "accomplishment auditors." The function of the 86 new employees is related to accountability theory, which maintains that schools should be held accountable for the successes and failures of their students. This an obvious challenge to education.

Industry, however, has been the one to leap to this challenge; the accountability theory is presently being tested on a large scale in the twin cities of Texarkana on the Arkansas-Texas border. Different firms involved in educational technology were invited to bid for contracts, which have as their aim raising the reading and math skills of potential high school dropouts. The company that won the contract was Dorsett Education Systems of Norman, Oklahoma.

Dorsett claims that in eight weeks, given students who lag two or three grades behind, it can successfully raise their performance by one grade level, as measured by achievement tests. If its goal is met, the company will be paid $80 per student; if it fails, the company will have to pay a cash penalty. If the required results are accomplished in less than eight weeks, a cash bonus will be awarded the company. The testing will be done by an independent project manager hired by the school systems. The project, which is being funded by the U.S. Office of Education,

is slated to run five years and will cost $3 million.

Open Court, a textbook publisher in La Salle, Illinois, has also put its product "on the line." It now guarantees to teach first graders to read at grade level, and promises reimbursement for the program materials if they fail.

In addition, part of the pressure to "show results" has come from the federal government in an indirect manner. The N.A.B. (National Alliance of Businessmen), in an effort to respond to the government's pressure to hire the "hardcore," had developed a "hire-fire" concept. Previously unhireable individuals are first placed on the payroll and then trained. For the business to protect its investment, the training program must work. And the government (since it is contributing an average of about $2,800 per trainee) is anxious to see empirical evidence of its success. The educators (in this case, the various business concerns), are being held accountable.

At the same time, with the growth of public concern for academic achievement, school systems are taking steps to provide some sort of concrete evidence of their success. In Columbus, Ohio, aptitude tests were given to sixth and eighth grades and the median results compared with national norms. Then, in a unique move (resulting from community pressure), the results were made public.

Also, the state legislature of Michigan has authorized standardized testing in reading, English, and math to be administered to fourth and seventh graders in every public school in the state. The results will be coupled with statistics relevant to the socio-economic status of the geographic area where the tests were given. These too will be made public.

The problem with the Columbus and Michigan projects, however, is that the public has forced the use of figures which are virtually meaningless. To pit any group of students against a national norm is to tie education to an arbitrary standard. These statistics will only tend to bring the various school systems toward a middle, rather than leading them to a true evaluation of whether learning has taken place. To be truly meaningful, the measurement devices employed must be related to the objectives as determined by the instructors and institutions. It is only because devices (called "criterion reference" tests) have not been developed to any great extent that school systems and state legislatures have turned to the only available numerical scores, the essentially irrelevant national norms.

However, if these experiments are successful, that is, if the results satisfy the public, we can look forward to industry taking over the teaching process in other areas. Education is big business. Nearly one-third of our nation is engaged, full-time, in the educational process. The American educational establishment costs $64.7 billions. It is one of the nation's largest growing enterprises. In 1969-70, higher education alone cost $22.7 billions.

John Roueche and John Boggs pointed out another consequence of these facts:

> Because of the increased need for funds, boards of trustees, parents, efficiency minded legislators, and the public are asking whether institutions are getting the maximum value from each dollar expended (12).

The implications are that if schools and colleges do not take the initiative, holding themselves accountable for improvement in instruction, for finding better ways to document learning, and for resolving the problem involved in education itself, others will take over the process for them.

A DEFINITION OF TEACHING

Based on a 1966 survey of 1,250 colleges by the American Council on Education, Alexander Astin and Calvin Lee (2) reported that most institutions claim that teaching effectiveness is a major factor in determining a faculty member's value to the institution. The community college is often looked upon as a "teaching institution." Although everyone agrees that "teaching" is the most important function of education, agreement on what constitutes "good teaching" has not been achieved.

Benjamin Bloom (3) has written that education exists for the purpose of providing experiences that bring about desired changes in the thoughts, feelings, and actions of students. In this light, Cohen and Brawer (4) have said:

> The only valid and stable measure of effectiveness is pupil change--simultaneously the end product and the single, operationally measurable kind of criterion that can describe teaching effectiveness.

Hence, indices of student change in desired behavior, operationally defined from educational objectives, may be the best way to measure teaching effectiveness.

THE PROCESS OF EVALUATION

In evaluation we are involved in making a judgment about something. To make this judgment, we choose some observable event that we infer represents a demonstration of what we are interested in judging. We then take measures of this observable event and compare these measures to a standard we have set up about it.

The thing about which we want to make judgments we call the "criterion referent." In education our criterion referent is teaching effectiveness. The observable event, which we choose to represent the demonstration of what we are interested in judging, is called the "criterion measure."

The usefulness of a criterion measure is usually determined by the degree to which it measures what it claims to measure. Its quality is judged relevant to the amount of bias it contains.

JUDGING CRITERION MEASURES

To be able to evaluate the usefulness and quality of an evaluation scheme, we must be able to evaluate its criterion measures. Robert L. Thorndike (13) classifies criteria as ultimate, intermediate, and immediate. The ultimate criterion is the one that is most relevant, but is very difficult to obtain. The criterion behavior, or observable event, is frequently very hard to measure because it is usually so intertwined with extraneous or uncontrollable variables that we cannot use it. In terms of teacher effectiveness, the teacher's long-term impact on changing a student's behavior is the ultimate criterion. This is the product of teacher effectiveness.

The intermediate criterion is one step removed from the ultimate criterion. It is used when the ultimate criterion cannot be. An example of this is practice-teaching ratings. We assume that a new teacher who was rated highly in practice teaching will be an effective teacher. We support our decision that a relationship exists by correlating practice-teacher ratings with supervisor's ratings after some teaching has taken place. We have increased the possibility for error, however, when we have used an intermediate criterion. In our example, the correlation may reflect the student teacher's ability to "brown nose" rather than his teaching effectiveness. A relationship like this one cannot be used to support cause and effect.

The immediate criterion is usually the most accessible and the least useful. If teaching effectiveness is the ultimate criterion, then the immediate criteria might be such things as a faculty member's holding a Ph.D., his years of experience, the amount he has published, or some other fact; but, as a criterion, it is at least two steps removed from what we are really interested in. As one goes from the ultimate to the immediate criterion, convenience and accessibility increase as relevance and importance decrease. We are dealing with the immediate, or at best, intermediate criterion to the extent that we simply describe a faculty member and his credentials. We are dealing with the ultimate criterion to the extent that we can measure changes in the students' behavior (on relevant variables). As Gustad states: "If we are in a position to use an ultimate criterion, we can afford to abandon the others" (7).

SOME COMMENTS ON PRESENT EVALUATIVE METHODS

Most current evaluation schemes are unsatisfactory ways to evaluate teaching effectiveness as compared with what we could achieve by developing the alternative methods available to us. Current schemes measure the process of teaching rather than its product. It is not so much that they are useless. They are legitimate ways to investigate the teaching process itself; i.e. behavior that a person exhibits while he is teaching or the characteristics of people who teach. (In fact, this type of investigation should lead to useful prediction schemes for future teachers.) But it is because of what they measure and how unreliably they measure it that they cannot be used as a basis for the purpose of making fair, valid, and meaningful judgments about teacher effectiveness (particularly as these judgments relate to individual personnel decisions; i.e., dismissal, promotion, salary increases, etc.). Administrators and institutions reward and honor teachers for supplying and propagating measures of immediate criteria, rather than identifying the ultimate criteria and rewarding faculty who meet them. Attention, therefore, should be given to alternative methods that deal directly with the teacher's impact on his students.

SOME ALTERNATIVE APPROACHES

Based largely on Douglas McGregor's performance-analysis method, which was designed for industry, a scheme has been proposed for evaluating teachers by student attainment (9). This scheme is based on the premise that, before educational procedures can be established and teacher effectiveness assessed, the ends of instruction must be agreed on. The essence of this scheme is the development of a carefully selected set of behavioral objectives for the student to accomplish and an assessment of the skills, attitudes,

and uses of knowledge exhibited by the teacher. These objectives would be developed cooperatively by the teacher and the administrator. It is believed that a necessary factor is mutual agreement between teacher and administrator on what would be accepted as evidence of student attainment of the specified objectives. (The teacher would be allowed to use any method of teaching he feels is best to achieve the objectives.)

Through advance agreement on the objectives to be achieved and the evidence that will be accepted that the teacher has been successful in changing the behavior of students, a shift from judging according to procedures followed (process) to judging according to the results produced in students (product) would be achieved. Also, by using a pre-test, the previous achievement level of students is taken into consideration.

While it is an improvement over current methods, certain aspects of this approach can be criticized. Popham (11) says that there is a difficulty involved in developing behaviorally based pre-tests and post-tests sufficiently reliable and discriminating to serve the purpose of teacher evaluation. Indeed, Gustad maintains that the development of adequate devices for measuring student progress toward course objectives would be "one great step forward" that could be taken immediately. He asserts, "It can be said that the teacher's examinations are an ultimate criterion since the teacher is the one who establishes the goals." He warns, however, that to feel comfortable with this, teachers' tests stand in need of great improvement (7).

J. Myron Atkins (1) also criticizes this approach (measurement by obtaining objectives). He claims that evaluation of specified content does not provide for outcomes expected as an outgrowth of many courses. Another difficulty, he says, is that this approach focuses on rather short-term behavioral changes and tends to obscure the long-term goals. He warns that stating objectives too early may obscure potential significant outcomes that do not become apparent until later because they are seldom anticipated.

A RESOLUTION

Work is presently being done on expanding the theory and technology of measurement to accommodate the problems in criterion test (mastery of objectives) instruments. In-service training programs for teachers could involve learning how to construct and improve their pre- and post-tests. (We already have ways to assess higher level cognitive behavior.)

Concerning Atkin's criticism, Bloom has already indicated that it takes some time before relevant and variable objectives and learning procedures can be set up. He says it may take as many as three or more attempts before the practice is demonstrable (3). It must also be recognized that most faculty members have never had training in writing course objectives or developing good test instruments. Training and time are, therefore, necessary to make this evaluation scheme viable, but the benefits for doing so far outweigh the time and costs involved.

Bloom has recognized that some objectives require learning experiences simultaneously in several parts of the curriculum if growth is to be adequately reinforced and that significant growth in certain objectives may require a sequence of learning experiences over several semesters.

We submit that these problems can be resolved if critics will recognize that students encounter more than one teacher when they attend our colleges and that the total college environment and its support services directly or indirectly influence the students' learning. We further submit that, while the teacher should be held accountable for teaching his course objectives, he alone cannot be held accountable for objectives that require several other teachers and courses to develop. We can, however, hold a department or a curriculum division accountable for measuring and documenting its effectiveness in obtaining these objectives. Furthermore, we can hold a college accountable for the attainment of its educational goals and objectives, which may require the 'several semesters' for the student to develop.

This would kill two birds with one stone. We would have an effective evaluation scheme to measure not only faculty members, but administrators as well. Contemplate what this might do for "causing learning."

Furthermore, data in the form of "hard copy" would be readily available to those elements of the community who are demanding evidence of an institution's effectiveness. The pyramid would move from the evaluation of the college as a whole through measures of the effectiveness of various curricula and further measures of departmental effectiveness, to measuring the foundation of any institution--the effectiveness of its teachers.

WHERE TO BEGIN

The problem, then, is not what to do, but how to do it. Though many teachers and administrators admire the goal, they would be hard pressed to find a starting point.

The first step, as always, must be commitment--from the individual instructor to the local board. The school's intention to pursue an approach oriented toward developing measures of its effectiveness and based on sound bahaviorist principles must be felt throughout the institution. Once this is achieved, faculty volunteers could be called on to participate in the development of pre- and post-tests to measure their effectiveness. Merit pay might be used initially to reward those who achieve success over and above the standard.

After the successful implementation of such a program, which would of necessity force the development and refinement of course and program objectives, the determination of curricula objectives would follow. Teams should then move "into the field" to define, in behavioral terms, the skills necessary for success in various occupations and professions. Community colleges would thus truly answer the needs of the communities they serve and could demonstrate to the public that they are doing so.

Finally, the assembled objectives of the various curricula, coupled with the objectives of the institution (many of which now fall into the category commonly called "general education"), would represent the objectives of the college community as a whole, and measures of these could be developed. (For example, follow-up studies might be done to determine how many graduates became regular voters.)

There will be problems, of course, as objectives must be continually subject to close scrutiny and revision. It might be years before a completely operational system could be developed. Eventually, administrators might find themselves in roles more relevant to the education process and less related to "paper work." Real leadership, which could be concretely measured, would be the end result.

Ultimately, a measure of institutional effectiveness would result. An accurate mode of comparison would be achieved and complete accountability, open and public, would arrive.

At John Tyler Community College a first step has been made toward developing such a system. The groundwork has been laid for a method of evaluating faculty on the basis of their students' attainment of objectives. Working with the staff of the Regional Education Laboratory for the Carolinas and Virginia, we have conducted in-service programs for our faculty and have developed methods and materials for individualizing instruction. The local board has endorsed a total commitment to the systems approach to teaching, and new faculty are hired with the understanding that they will be held accountable for student learning. Furthermore, a committee of faculty and students has been working on the development of objective criteria for evaluating administrators as well as faculty. Thus, the early steps have been taken. Much more work is necessary before a truly viable program emerges, but the beginnings are under way. It is hoped that others will assist us in the work, for it is only through such cooperation that we can succeed, and time is short. William A. White has said that, "Unless the free

are brave, they will no longer be free." Today, faculty and administrators are free to choose and develop the means that will be used to evaluate them. To begin this task takes a certain amount of bravery. If we lack that bravery, then, most surely, we will lose that freedom.

REFERENCES

1. Atkin, J. "Some Evaluation Problems In a Course Content Improvement Project", Journal in Research in Science Teaching, 1: 129-132, 1963

2. Astin, Alexander and Calvin B. T. Lee, "Current Practices in the Evaluation and Training of College Teachers," Improving College Teaching, American Council on Education, Washington, D.C., 1967

3. Bloom, Benjamin S., "Testing Cognitive Ability and Achievement," Handbook of Research on Teaching, ed. N. L. Gage (Chicago, Ill: American Educational Research Association, 1963)

4. Cohen, Arthur M. and Florence B. Brawer, Measuring Faculty Performance. Monograph No. 4, American Association of Junior Colleges, Washington, D. C., 1969 (ED 031 222)

5. Dressel, Paul L., "The Current Status of Research on College and University Teaching," The Appraisal of Teaching in Large Universities, ed. W. J. McKeachie, University of Michigan, Ann Arbor, 1959

6. Dressel, Paul L. and Associates, Evaluation in Higher Education. Houghton Mifflin Company, Boston, Mass., 1961

7. Gustad, John W., "Evaluation of Teaching Performance: Issues and Possibilities," Improving College Teaching, American Council on Education, Washington, D.C., 1967

8. Israel, Jack W., "Innovation in Evaluation: Teacher Assessment by Objectives," (Seminar Paper). Los Angeles, University of California, 1969 (ED 029 625)

9. McGregor, Douglas, "An Uneasy Look at Performance Appraisals," Harvard Business Review, 35; No. 3 (May-June, 1957)

10. McNeil, John D., "Antidote to a School Scandal," The Educational Forum, November 1966, 31 (1), 69-77

11. Popham, W. James, "The Performance Test: A New Approach to the Assessment of Teaching Proficiency," The Journal of Teacher Education, Vol. 19, No. 2 (Summer, 1968), pp. 216-222

12. Roueche, John E. and John R. Boggs, "Junior College Institutional Research: The State of the Art," Monograph No. 2, American Association of Junior Colleges, Washington, D.C., 1968

13. Thorndike, Robert L., Personnel Selection, John Wiley and Sons Co., New York, 1949

14. Tyler, Ralph W., "The Evaluation of Teaching," Preparing College Teachers, ed. A. D. Albright and John E. Barrows, University of Kentucky, Lexington, Kentucky, and Southern Regional Education Board, Atlanta, Georgia, 1959